# The Faults and Foibles of Energy Dispersive Spectrometers 2016 edition

Nicholas W. M. Ritchie National Institute of Standards and Technology Gaithersburg, MD 20899 nicholas.ritchie@nist.gov

# Imagine the perfect detector

- Infinite resolution
- Quantum efficiency of unity for all energies
  - One x-ray in → one measured x-ray
- Infinite throughput
- No artifacts
- 4  $\pi$  collection solid angle (with position sensitivity)

The remainder of this talk will discuss the consequences of the various ways these ideals are violated and what to do about it.

# Overview

- Part 1 Hardware
  - Detectors
  - Electronics
  - Processing
- Windows
  - UTW / Si-nitride
- Artifacts
  - Escape peaks
  - Edges
  - Coincidence events
  - Resolution
  - Decreasing efficiency at high E

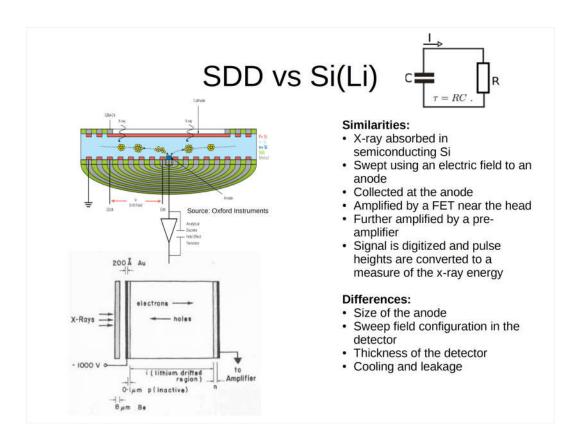
- Part 2 Hygiene
  - Detector orientation
  - Optimal WD
  - Pulse process time
  - Selecting a probe current
  - Laboratory QC
  - Detector / probe current linearity

To understand how best to operate an energy dispersive spectrometer, requires an understanding how the detector converts x-rays into a spectrum.

The limitations and foibles of the detector technology directly play into the choice one makes to select the optimal measurement parameters including working distance, process time, probe current etc.

You can learn rote procedures to determine the optimal parameters but the knowledge is brittle and won't help you to diagnose problems with things go wrong.

This presentation takes the perspective that understanding the underlying detector technologies will help you to make better measurements. It will pass back and forth between detector design and the consequences



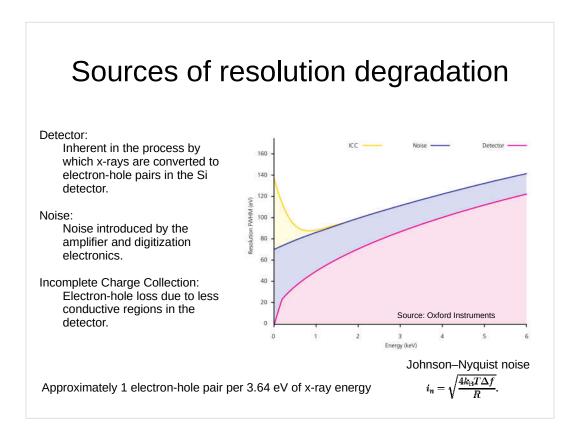
The basic detection mode is identical between the SDD and the Si(Li) detector.

As X-rays pass through the active detector region, they can be absorbed by the Si producing an ionized Si atom and an energetic electron. The ionized Si atom decays producing 95% of the time an Auger electron and 5% of the time a Si K x-ray. The energetic electron and the Auger electron loose energy to the Si through production of conduction electron-hole pairs. Each conduction electron-hole pair takes about 4 eV to produce.

The electrons are swept by electric fields to the anode where the resulting current is measured on an x-ray event-by-event basis.

The biggest difference between an SDD and a Si(Li) detector is capacitance. SDD detectors have tiny anodes with tiny capacitance while Si(Li) detectors have large anodes with large capacitance. As with an RC circuit, high capacitance (and equivalent resistance) lead to large time constants (τ) or equivalently to slow detectors.

The fact that SDD can be cooled with solid state Peltier coolers rather than liquid nitrogen is further reason to appreciate this innovation.

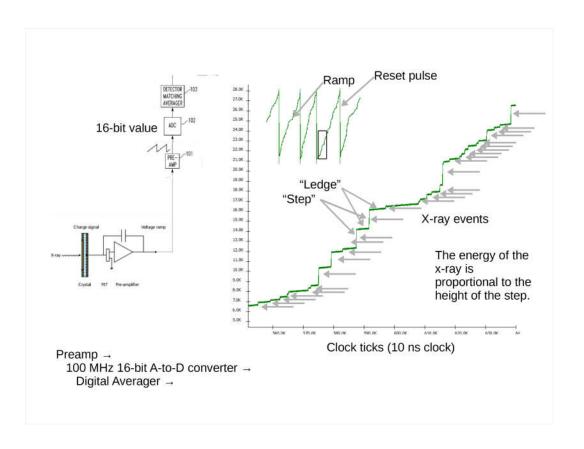


There are three dominant sources of resolution degradation in a well designed SDD system.

Detector noise is fundamental to the mechanism by which xrays are converted to signal. We are stuck with this unless there is a radical change in technologies (like microcalorimeters.)

Electronic noise – Well designed electronics can minimize this contribution but never eliminate it. Fundamental solid state physics like Johnson-Nyquist noise limit our ability to amplify signals at high speed. The best detectors/pulse processors are approaching theoretical limits.

Incomplete charge collection occurs when generated conduction electron-hole pairs don't reach the anode. This usually occurs when the x-ray is absorbed close to the front surface of the detector (as is often the case for low energy x-rays.) Modern detector design have significantly reduced the thickness of the "dead layer" on the surface of the detector and as a result significantly reduced incomplete charge collection.

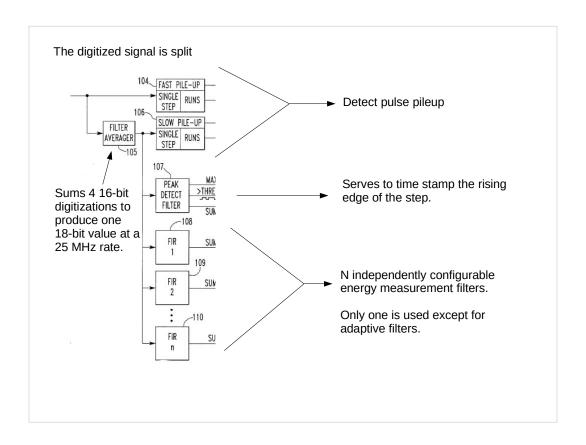


All modern pulse processors convert the output current from the SDD into a digitized value as soon as possible. This requires a fast (100 MHz or so) analog-to-digital converter (ADC) with a high bit depth (16-bit or so) to convert the ramp signal from the detector into accurate measurements of the step height.

The digitized signal can then be further processed without additional degradation using digital algorithms implement in an FPGA or equivalent fast electronic processing device.

The output of the detector is a ramp signal with two distinctive features – ledges and steps. The ledges represent time between x-ray events and the steps represent x-ray events. The height of the step corresponds to the energy of the x-ray.

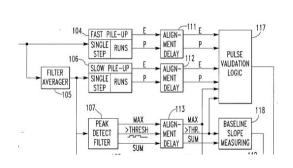
There are also reset events to keep the size ramp within the range of the ADC. The reset corresponds to zeroing the voltage across the current amplifier capacitor. Reset events are a source of dead-time.



The ADC value is split among various processing schemes to detect pulse-pileup, to measure the timing of the event (independent of amplitude) and the amplitude of the event.

Because the value is digital all of these processing schemes can occur simultaneously ("in parallel.") In fact, it is possible to apply multiple different amplitude measuring filters simultaneously and report only the one that produces the most accurate results given the input data. (This is how "adaptive time constants" work.)

Time stamping the x-ray is critical for x-ray spectrum imaging in which accurate knowledge of the beam scan coordinates at which the x-ray was generated is critical for accurately reconstructing the map. ("Position tagged spectrometry")

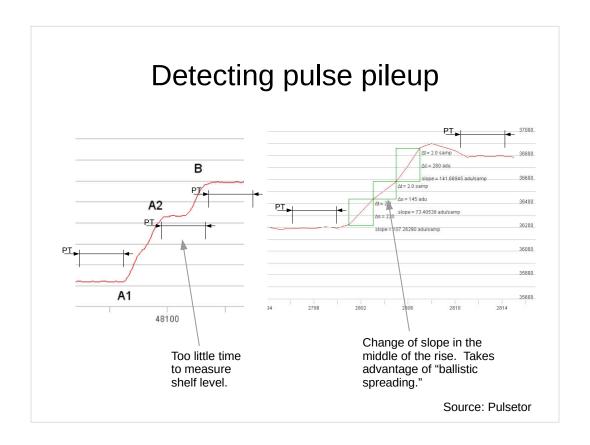


Pulse validation takes as input:

- Fast pile-upSlow pile-upPeak detect

Pulse validation does not take:

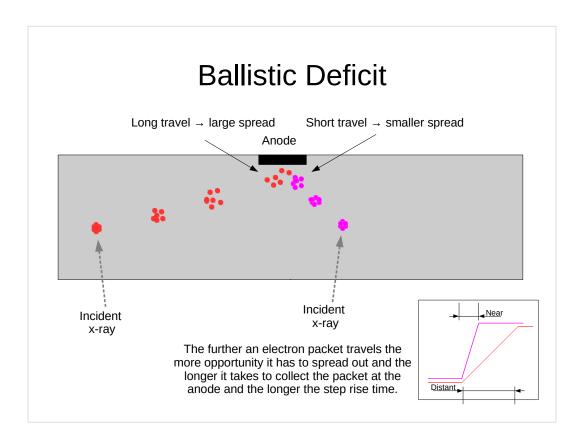
· Measurement filters



Pulse pileup is detected by

- 1) looking for shelves that are too short to measure.
- 2) looking for changes in the slope of a step.

The second mechanism works because of "ballistic deficit." In short ballistic deficit occurs because the bundle of electrons created by each x-ray travel different distances to the anode. While the bundle travels the initially tight bundle expands.



A electron packet produced far from the anode takes longer to reach the anode, expands more and produces a step with a longer rise time.

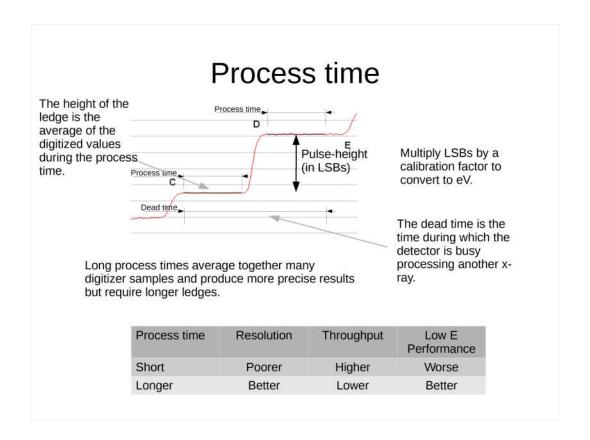
An electron packet produced closer to the anode spreads less and produces a shorter rise time.

A change in slope during a step is indicative of two pulses producing the step.

Pile rejection is much easier at high energies and at low count rates. At low energies (small steps), it can be difficult to differentiate random noise from a step. At low energies, it can be difficult to determine whether the slope has changed.

To get the best efficiency at low energies takes long pulse-process times and low count rates to minimize the likelihood of pulse-pileup.

There is also delay inserted between the measurements of the leading and trailing ledge heights. This delay is long enough to handle the longest rise time.

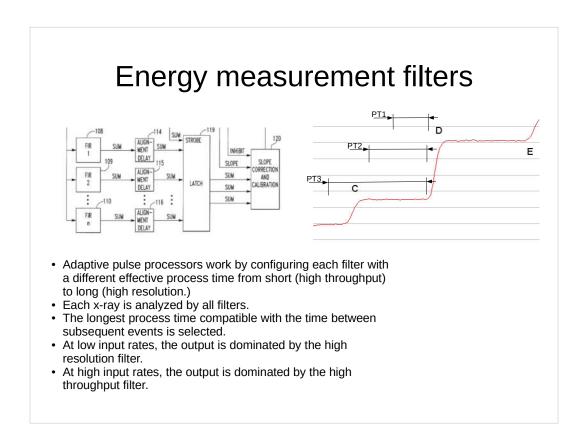


Averaging together multiple digitized ADC values is the key to enhancing resolution. The longer the averaging the more accurate the measured height of the ledge (up to a point of diminishing returns.) However, the length of the ledges must be longer (lower input count rate) to support longer integration times.

Since x-ray events arrive at random times, there will always be events that occur to close together.

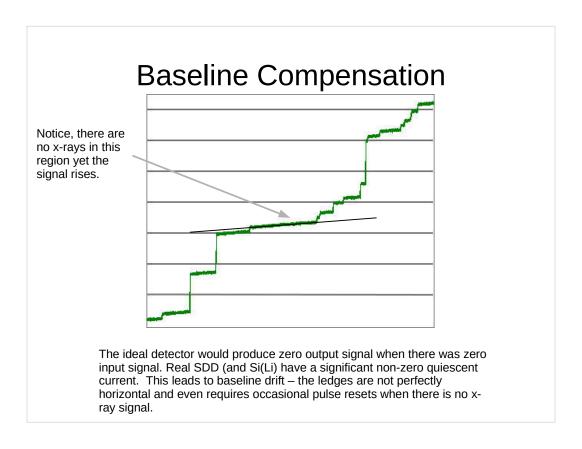
In general, shorter process times lead to poorer resolution, higher throughput and worse low energy performance.

Longer process times lead to higher resolution, lower throughput and better low energy performance.

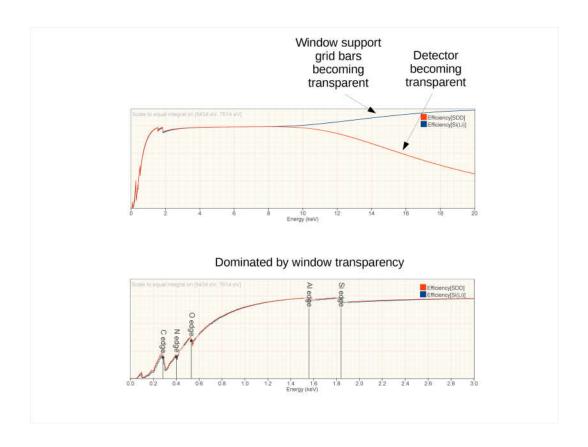


Multiple energy filters can be programmed for different process times. The longest process time compatible with the current ledge can be selected for each x-ray. The result balances high resolution when x-rays are well spaced and high throughput for crowded x-rays.

However, adaptive pulse processing schemes show changes in resolution with throughput. At low throughput, the long filters dominate and the spectrum is high resolution. At high throughput, the short filters dominate and the spectrum is lower resolution.



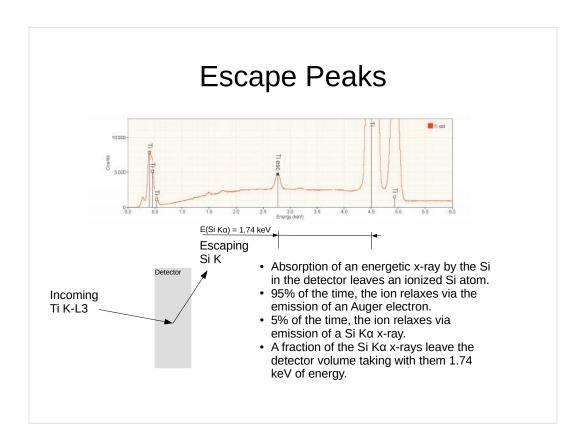
SDD also show significant leakage current (current on the anode with zero input signal.) This leakage current must be compensated for the most accurate measurements. The compensation operates by measuring the slope of the ledges and correcting the measured energies for this slope.



SDD are thinner (~0.50 mm) than Si(Li) detectors (3-5 mm). This results in a decrease in efficiency starting around 10 keV for SDD relative to Si(Li).

Coincidentally, the thickness of the window supports are approximately the same thickness as an SDD so just when the window grid bars are becoming transparent (leading to an increase in efficiency in Si(Li) detectors from 80% towards 100%), the SDD start to decrease in efficiency.

At low energies, the efficiency of a Si(Li) or SDD is largely limited by window transparency which in polymer windows is dominated by the C, N, O in the window.



Escape peaks happen when the 5% of Si ionizations in the detection region that lead to Si K x-rays escape from the detector. The associated energy (E=1.74 keV) is lost to the system producing an artifact peak at 1.74 keV below the original peak.

Fun fact: The artifact peak has the resolution you'd expect for a peak at the escape peak energy (not the resolution of the source x-ray.)

# **Detector Checks**

- Take off angle
- Process time
- Optimal working distance
- Optimal probe current
- Throughput linearity
- Energy calibration linearity
- Setting up a simple QC program

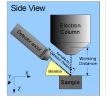
Modern EDS detectors are amazing spectroscopy tools but the quality of the results depends upon the detector being configured and operating correctly. There are a number of simple tests that you can perform to verify that your detector is performing optimally.

Some of these checks can be performed occasionally when a detector is first installed or a major change has been made to the instrument or detector.

Other quick checks should be performed very time that you sit down at the instrument to demonstrate that your detector continues to perform its best.

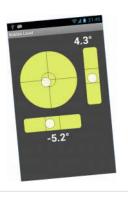
# Checking the detector elevation

 The correct detector elevation is critical for good quantitative analysis.





- · Too often it is recorded incorrectly.
- Examine your vendors spectra files (is the elevation angle correct?)
- You may be able to use a cell phone app to verify the elevation angle.



To perform good quantitative analysis, your quantitative software needs to know the elevation angle at which the detector is mounted in the chamber. (Elevation angle is sometimes called "take-off angle." I differentiate the two. Elevation angle is a fixed property of the instrument/detector system whereas take-off angle can vary with sample tilt. When the sample is normal (perpendicular) to the beam, the take-off angle equals the elevation angle.) Typical elevation angles are between 30° and 50° with the most common values being 35° and 40°.

You can measure the take-off angle with a protractor or using the tilt-sensors in a typical smart phone. (Make sure to calibrate your phone. A 30°-60°-90° triangle makes an excellent test sample.)

# Choosing the process\* time

- The variation of resolution with respect to process time is less extreme with an SDD
- For quantitative analysis, throughput (almost always) trumps resolution.
- Throughput is usually limited by coincidence events rather than process time.
- The relationship between process time and pulsepileup depends upon subtle settings within the pulse processor (probably outside of your control.)

\* "Process time" aka "shaping time" aka "resolution setting" aka "throughput setting"

Process time determines throughput, resolution and low energy sensitivity.

- Resolution is important but it is rarely worth the loss of throughput associated with getting the highest resolutions. Moderate resolutions usually produce better (more precise/accurate quant results) than the highest resolutions.
- Throughput (as defined as output count rate divided by input count rate (OCR/ICR)) is maximized at short process times.
- Typically however, the maximal usable throughput is not limited by process time but coincidence event rates (pulse pileup) instead.

# **Process Time**

- A moderate process time is probably the best place to start
  - High resolution modes (long process time) tend to hurt throughput while only providing moderate improvements in resolution.
  - High throughput modes (short process time) are limited to much less than their maximum rates by coincidence events
  - Moderate resolution balances resolution, throughput and coincidence events. Aim for within about 5 eV of the best resolution.

### Considerations

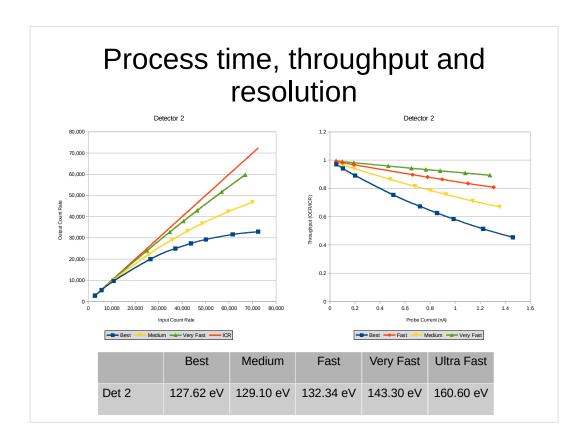
- Quantification is optimized by high throughput with low coincidence events (resolution really doesn't matter that much.)
- Low energy x-rays (Be, B, Ca L, Ba M etc) are likely to be more visible at the longest process time because they involve differentiating very subtle distinctions between noise and steps.

Moderate process times usually provide the best compromise between throughput, resolution and low energy sensitivity.

You typically can't take advantage of fast throughput / poor resolutions settings because of pulse pileup / coincidence events.

High resolution modes typically come at the cost of a significant loss of throughput.

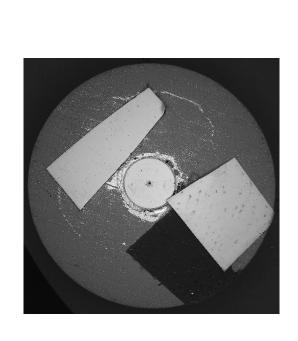
For spectroscopy of low Z materials or other low energy x-rays, you may need to optimize your process time to detect these x-rays. This usually involves a longer process time.



These plots show the relationship between the input count rate (input counts per real-time second, proportional to probe current) and the output count rate (measured x-rays per real time second).

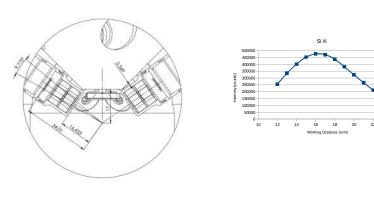
The ratio of these two numbers is the throughput which is unity when every incoming x-ray is measured.

The resolutions specified are for my detector using various different pulse process times from long ("best") to short ("ultra fast").



# The "Optimal Working Distance"

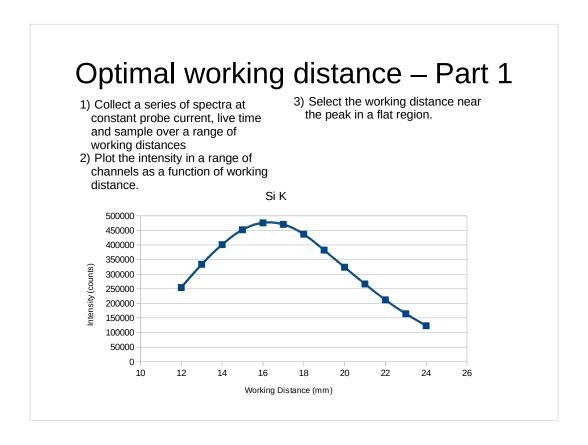
- Working distance at which the take-off angle equals the nominal take off angle?
- The working distance at which the axis of the detector intersects with the optic axis?
- The working distance producing the largest x-ray flux?
- The working distance at which the flattest part of the field is centered?



The "optimal working distance" is the lens-to-sample distance ("working distance") for which the EDS detector alignment is optimized.

Often you can get schematics from your SEM or EDS vendor that shows the design optimal working distance. This should be your starting point but you should verify that your instrument performs as designed.

Working at the optimal working distance is important because this sample distance is the place on the throughput vs working distance curve with zero slope ie. it has the smallest sensitivity with respect to slight errors in sample position.



Determining the optimal working distance involves measuring the throughput at constant current / constant live-time over a range of working distances. Any sample will do but Cu or Al are good choices. (Look at the net intensity in the K lines.)

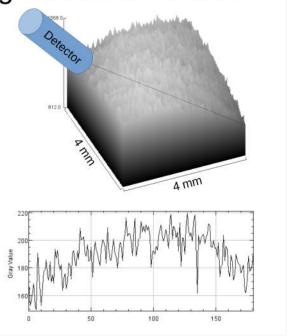
Plot the net intensity vs working distance. Select the working distance that produces the largest net intensity.



- Mount a flat smooth piece of Al in the SEM at the optimal working distance for x-ray acquistion.
- Acquire a x-ray spectrum image of as broad a field-of-view as the SEM is capable of imaging.
- Export the spectrum image as RAW file
- Process the spectrum image using FIJI (ImageJ) to "Plot Profile" and "Surface Plot" the variation in AI intensity with position.

### Verify:

- The sweet spot of the detector is in the center of the field
- The size of the sweet-spot (± 1% in intensity)



The ideal x-ray detector is oriented such that the "sweet spot" - the most sensitive Ideally, the x-ray detector orientation should be optimized for the intersection between the electron beam axis

# Processing a spectrum image with FIJI Export the spectrum image data in "RAW" format. Import the data using "Import → Raw" to create a "stack" Identify the band of channels in the stack representing "AI K" "Reslice" the "stack" until you viewing it from the image plane Perform an "Image → Stack → Z-project" to extract a plane representing the "AI K" data Use "Analyze → Plot Profile" or "Analyze → Plot Surface" to visualize the intensity profile

Most x-ray systems collect x-ray spectrum image data in a native streaming "position sensitive" format. Most systems also offer a utility to convert the native streaming format into a generic "raw" binary format which consists of a contiguous array of binary encoded spectra.

X-ray spectrum image data in "raw" can be read into and processed in ImageJ using the "stack" tools.

The stack is likely to be misoriented to use the "Z-project" tool so it is usually necessary to use the reslice tool to pivot the data in memory.

Use the "z-project" tool to extract the channels corresponding to the K line.

Use "Analyze – Plot profile" or "Analyze – Plot Surface" tools to visualize intensity profile.

# Determining the "optimal probe current"

- The optimal probe current produces a large number of measured x-rays with a manageable number of coincidence events.
- Dead time is no longer a good way to select the optimal probe current.
  - Dead time varies too much from vendor-to-vendor.
- You should determine (and maintain) an optimal probe current for each beam energy at which you work.
- There may be situations (like beam sensitive samples or trace element analysis) in which it is necessary to use less than the optimal probe current.
  - Consistency is the key to reliable quantitative EDS. While it is possible to use a range of different probe currents and compensate in the math, it is more reliable and less susceptible to the unanticipated to maintain a consistent probe current.

Now that we know that the detector is oriented correctly and the optimal working distance, we can now determine the optimal probe current.

With Si(Li) detectors, the optimal probe current was usually determined by dead-time – a dead time of approximately 30% worked for most detectors in most circumstances. SDD detectors are different! The optimal probe current in an SDD is the current that

produces a large number of measured x-rays with a manageable number of coincidence events.

# A procedure for determining the "optimal probe current"

- Select a Al sample. Mount the sample at the nominal working distance and geometry. Configure the instrument for the desired beam energy.
  - Al is selected because it represents a near "worst case" an intense peak at relatively low energy
- Starting at a probe current producing a low output count rate and moderate coincidence events, measure a spectrum from the Al.
- Measure the probe current. Perform a background corrected integration over the Al K peak and the Al K+Al K sum peak. Compute, tabulate and plot the probe current and ratio I(Al K+Al K)/I(Al K).
- Increase the probe current by a factor of two each step and repeat the spectrum acquisition and analysis. Repeat until the sum peak is larger than acceptable.
- Perform a regression on the data to estimate the probe current that produces the ratio I(Al K+Al K)/I(Al K) meeting your limit of acceptability.
  - For routine analysis, a ratio of <1 % produces accurate quantitative analyses. For trace analysis, I may require 0.1 % or less depending upon the location of the sum peaks relative to primary characteristic peaks.

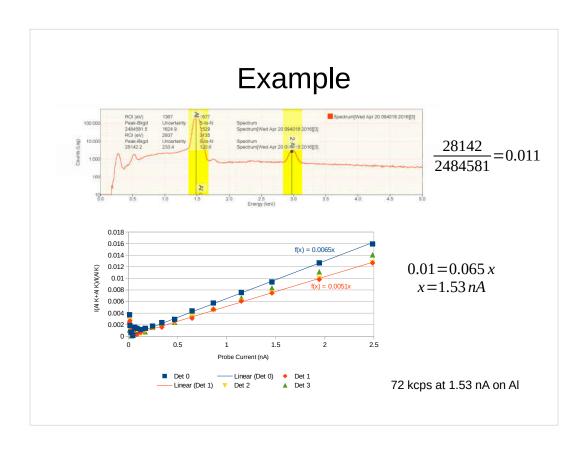
Al is a good sample for performing this optimization because it produces a large number of lowish energy x-rays and produces a significant coincidence peak.

You will need to determine the optimal probe current for each beam energy you intend to use.

Measure the intensity in the Al K and the Al K + Al K coincidence peaks at a range of different probe currents.

Plot the quantity I(Al K+Al K)/I(Al K) vs probe current. Determine an acceptable coincidence fraction and read off the plot the probe current that produces this coincidence fraction.

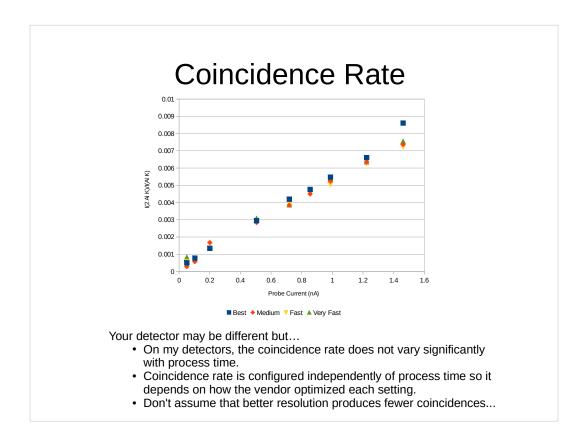
Use this current for routine analysis. There may be circumstances like trace element analysis which may require lower probe currents.



Using the peak integration tools in DTSA-II can be used to extract the intensities. Select a range of channels that covers the Al K and the Al K + Al K coincidence peak. A log scale on the ordinate axis helps to make both peaks visible. Use the "Integrate peak (background corrected)" tool to extract both region intensities which can be "Copy → Status Text" and pasted into a spreadsheet.

Plot the ratio vs probe current.

Use curve fitting or graphical methods to determine the maximum probe current that produces an acceptable coincidence fraction.



If you have time you can also measure the coincidence rate at various different process times.

# Linearity

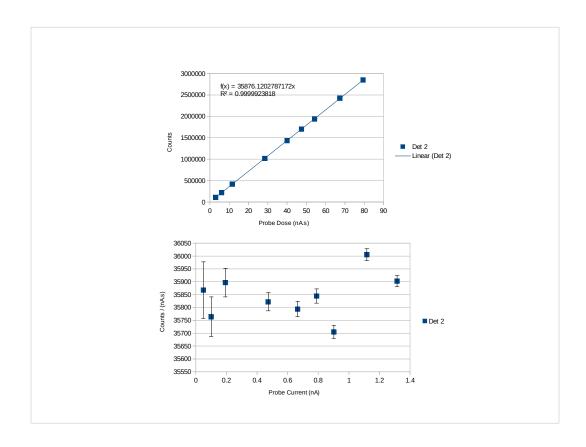
(Output counts)/(live time second) vs Probe Current

- Demonstrating linearity is easy except when it isn't...
  - It is easy if your detector and picoammeter are both linear. Otherwise...
- Select a material
- Use either an integrated probe current meter or a Faraday cup
- Collect spectra over a range of probe currents from pA to the maximum useable probe current
- Integrate the spectra over all energies
  - This partially compensates for coincidence events which move events from lower energies to higher but not for the fact that two x-rays become one recorded event
  - Plot "Probe Current (nA)" vs "Flux (Counts/(nA·s))"

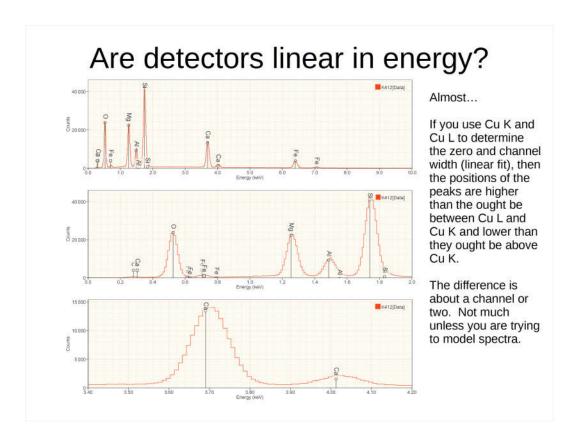
A 60 nA.s spectrum should contain the same number of counts (to within statistical uncertainties) regardless of whether the 60 nA.s is 10 nA for 6 seconds or 0.01 nA for 6,000 seconds where the times are "live times".

This test determines whether the pulse processor accounts for "dead time" correctly.

It also tests whether your probe current meter is linear with zero offset.



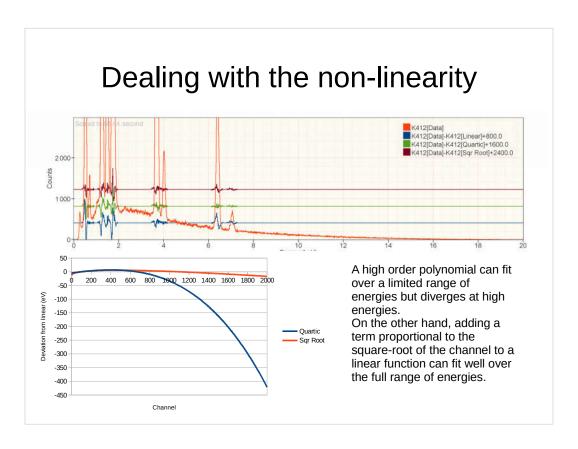
Plotting the net counts vs probe current is not a sensitive manner to determine linearity. It is better to plot (net counts)/(probe current) vs probe current. This plot should be a horizontal line.



We usually (meaning 99.99% of the time) assume that the correspondence between channel (energy bin) and energy is linear (E = A\*ch+B) and too a (very) good approximation this is true.

However, if you look carefully enough, this approximation breaks down. Depending upon how you calibrate your detector, the perceived errors will appear at low energies (below the low energy calibration point), at moderate energies (between the low and high energy calibration points) and/or at high energies (above the calibration points). (Of course, the definition of calibrated is that the low and high energy calibration points match perfectly.)

You may notice the KLM line markers don't perfectly align with the peaks (as shown here.)



Due to the magic of locally collected standard and reference spectra, usually there is no need to deal with the non-linearity. I've never seen it so extreme as to mess with quantification (fitting k-ratios).

However if one wants to model spectra as accurately as possible, it may be beneficial to consider alternative functional relationships between bin and energy. Two alternative models are higher order polynomials in bin or adding a term proportional to the square root of the bin.

Polynomials are sub-optimal because while they can fit very well within the calibration region, they tend to diverge outside the calibration region.

The functional form E = A\*ch+B+C\*sqrt(ch) seems to fit the actual non-linearity well with a minimal number of fit parameters and good behavior outside the fit region.

# **QC** Program

- Document that your detector was working correctly when you collected your data
- Identify problems with your hardware before they waste your time
- Determine when you need to recalibrate / restandardize
- Determine when your detector configuration has been changed (by other users / service personel)

QC is the way you demonstrate to your customers that you can be trusted yesterday, today and tomorrow.

- Having a long-term record of your detector performance enhances trust and eliminates worries that an current or historical data set was collected on a detector that can not be trusted.
- A good QC program and pre-established rules about re-calibration and restandardization will allow your laboratory to reuse standards and yet ensure sufficiently accurate results.
- A good QC program will also catch blunders due to changes in detector performance that can occur due to changes by other users or service personnel.

# My QC Program



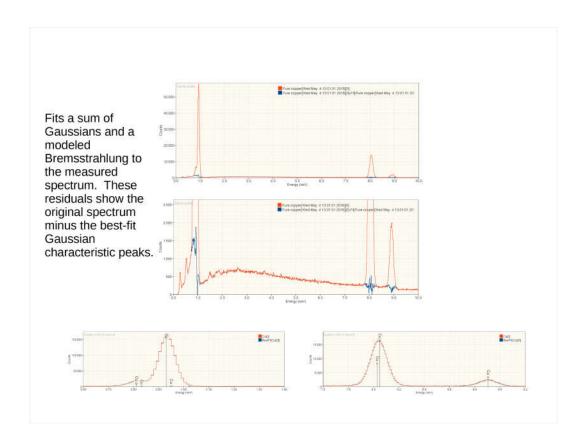
- Select 20 keV & ~780 pA probe current
- Measure PC
- Collect a 60 second Cu spectrum
- Use DTSA-II to perform a non-linear peak fit to the measured spectrum
- Extract intensities, calibrations, resolution etc from peak fit
- Record and tabulate

This protocol is implemented using DTSA-II's "Quality Control Alien".

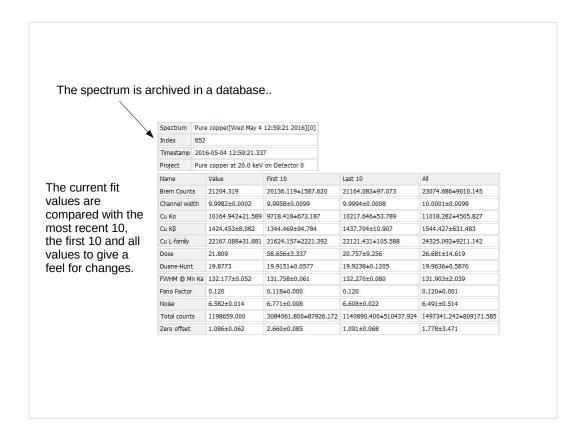
A QC program need not be complex. Collecting a single spectrum under consistent conditions and then extracting measurable quantities like line intensities, total number of counts, resolution, calibration and total counts, can be sufficient.

DTSA-II provides tools to perform implement a basic QC program using a single spectrum collected from a material of your choice at conditions of your choice.

DTSA-II can then produce "control charts" which allow you to compare today's performance with historical performance.

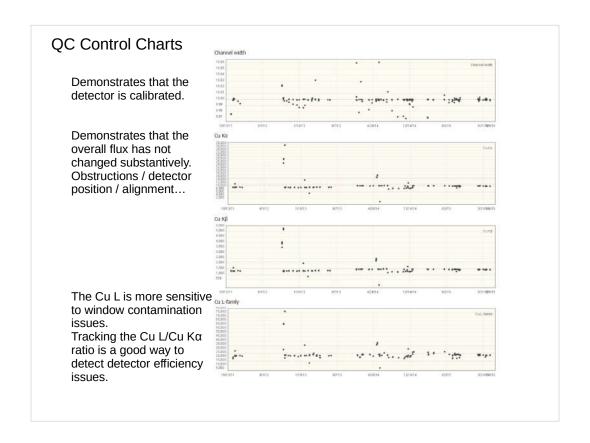


DTSA-II does a sophisticated non-linear fit of modeled continuum and peak shapes to the data. The result of these fits are highly reproducible and track many different aspects of the detector performance even when used on simple spectra (like pure Cu.)

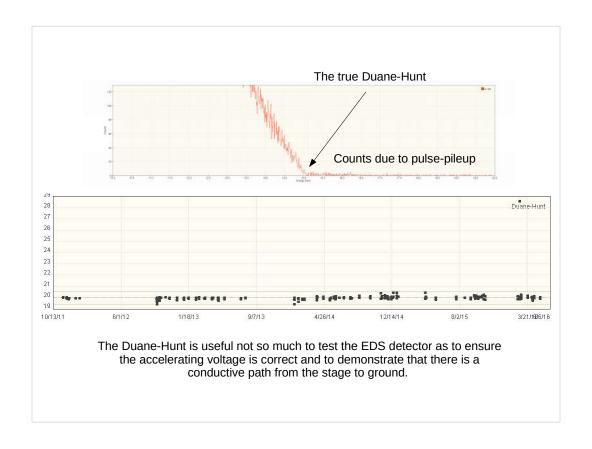


When a spectrum is added to the QC database, the current fit values are compared with recent (last 10), old (first 10) and all historical fit values. The result is reported as a table.

This allows you to quickly review the current values and determine if something is out-of-wack.



You can also generate an HTML report with control charts that summarize the historical values of each performance metric.



One instrument performance metric that can be extracted from the data is the true beam energy. The Duane-Hunt limit, the energy at which continuum production goes to zero, is a good measure of the actual energy at which the beam electrons strike the sample. Be careful not to mistake pulse-pileup which can occur well above the true Duane-Hunt limit for

continuum x-rays.

Nominally, all QC spectra should be collected at the same beam energy instrumental setting. It is rare that the actual beam energy diverges from the instrumental setting but not unheard of. If the emitted bias is wrong or if the stage ground has been disconnected this can be discerned in the Duane-Hunt limit.

# Choosing a detector for quantitative EDS

- #1 Linearity "Sine qua non"
  - The dead-time corrected output per nA·s should be constant with respect to probe current. (Make sure your picoammeter is also linear!)
- #2 Stability "Sine qua non"
  - Resolution and peak position stability
  - Almost all modern detectors are fine. Don't use adaptive process time!
- #3 Pulse-pair rejection at high throughput
  - Higher throughput with a small fraction of coincident events
- #4  $\Omega$  Solid angle
  - Collect more x-rays per nA·s
  - Detector area is a poor proxy for solid angle, solid angle is what matters!
- •
- #8 Resolution
  - Anything better than 135 eV is sufficient and almost every modern SDD is capable of much better than 135 eV.
- Finally, some thoughts on choosing a detector for quantitative EDS. The good news is that most vendors detectors are excellent and can produce excellent standards-based analyses (in DTSA-II and potentially also the vendor's software.)
- Detectors must have good dead-time correction leading to a linear response between dead-time corrected dose (live-time × probe current) and counts in the spectrum.
- The resolution and peak position must not change substantially with probe current.
- Good pulse pair rejection allows you to collect high count spectra in moderate real acquisition times.
- High solid angle allows you to collect high count spectra with moderate probe currents.
- High resolution at high throughput is nice but of secondary importance (so long as the resolution is "good enough.")

# Thanks!

Questions?

E-mail: nicholas.ritchie@nist.gov

Please feel free to contact me with questions.